PERGAMON Computers and Mathematics with Applications 39 (2000) 29–42
An International Journal computers & mathematics with applications
www.elsevier.nl/locate/camwa

Estimation Problems in an Input-and-Output System

F. S. MISTIRI
Faculty of Science, University of Garyounis
Benghazi, Libya

A. P. WANG
Department of Mathematics, Arizona State University
Tempe, AZ 85287-1804, U.S.A.

(Received June 1998; revised and accepted September 1999)

Abstract—We investigate smooth and filtered estimation problems in an input-and-output system. The Hamiltonian system is decomposed into internal and external states. With this approach, we are able to obtain general results which include some well-known ones. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords—Input-and-output system, Estimations, Scattering, Star-product.

1. INTRODUCTION

In an input-and-output system, inputs impinging on a black box are partly reflected, transmitted, and absorbed (or generated). Many physical processes can be viewed as input-and-output systems. The relations between inputs and outputs obey certain laws of physics for a given process. However, they are all governed by the same mathematical model called scattering theory. This unification enables us to study various physical processes with the same mathematical structure.

Physical processes, such as transmission line, cascade network, radiative transfer, dielectric wave propagation, neutron diffusion, etc., can all be considered as input-and-output systems; for more details, see [1–5]. The physics of these systems determines so-called coefficients or infinitesimal generators. Various processes have different types of coefficients. Once these coefficients are determined, all processes are governed by a set of identical mathematical structures, either in differential equation model or in algebraic model form. The differential equation model may involve Riccati type equations. That is why, for example, both the feedback control theory and the reflection operator for radiative transfer are involved in solving this nonlinear differential equation. The only difference between them are coefficients. In wave propagation in a dielectric medium, the coefficients are determined by medium parameters, such as dielectric constants and permeabilities [6,7]. In the analysis of specific intensities in radiative transfers, the coefficients are determined by medium parameters such as scattering coefficients, absorption coefficients, and

Please address all correspondence to A. P. Wang at 8428 S. Newberry Lane, Tempe, AZ 85284, U.S.A.

PII: S0898-1221(00)00127-9

phase functions [1]. For the stationary case, i.e., these parameters are constants in time, the coefficients related to wave propagation in dielectric medium and radiative transfers are "bounded operators". On the other hand, for the nonstationary case, the coefficients are unbounded operators [8,9]. This explains that well-known results in wave propagation in the dielectric medium can be translated to radiative transfer and *vice versa*. It also indicates the difficulties involved in the nonstationary case. The advantage of this model is that there are many well-known results in differential and operator equations which can be applied here. There are also results in particular on Riccati type equations. For related references in analyses, see [10] for the finite-dimensional case and references in this article for the infinite-dimensional case. For computational methods, see [11,12].

A simple algebraic model was obtained early by Stokes [13] in the study of reflection of a pile of identical glass. Slowly this model was extended to nonidentical glass and to the continuous case. After the full development of radiative transfer by Bellman and his associates [11], the neutron diffusion by Wing [5], and the transmission line theory by Redheffer [3], a unified algebraic model was established by Redheffer [3], so-called star-products. The algebraic structure is associative but not distributive under the star-products. Of course, it leads to differential equations by taking limits. The algebraic model explains, for example, the physics of multiscattering processes very clearly and is easy to understand. This model may be more suitable for digital computation [2]. It can even be used to solve inverse problems [14].

The purpose of this paper is the application of an input-and-output system to obtain some results in estimation theory. Estimation theory is seldom considered in this way because of the dimension of the process and the nature of representation used for the operators. After a short introduction of stochastic operators in Hilbert space, we state estimation problems in the frame of input-and-output systems. From the input-and-output point of view, internal and external states are naturally introduced. In fact, the decomposition of the Hamiltonian system into two states is mathematically necessary and sufficient. It turns out that the external state is the well-known Kalman filter. In the analysis, we extended the 2×2 scattering matrix to a 3×3 for convenience and necessity of our analysis. We obtain some new and old results. The most important is that we have established a relation between estimation theory and input-and-output systems.

2. STATE ESTIMATION

To investigate the estimation problem to be treated in this paper, we consider the following stochastic integral equation on some Hilbert space:

$$u(t) = u(a) + \int_a^t A(s)u(s)\,ds + \int_a^t \phi(s)\,dW(s), \qquad t \in [a,b], \tag{1}$$

where $A(s)$ is a deterministic linear bounded operator, $\phi(s)$ is a suitable transformation valued stochastic process, and $W(s)$ is a Hilbert space valued Wiener process.

To discuss the integral equation (1), we need some basic abstract probability theory. We consider throughout that $(\Omega, \mathcal{A}, \mu)$ is a probability space, where $\mathcal{A}$ is a Borel field and μ is a probability measure. We let H be a Hilbert space and $[a,t]$ to be a finite interval.

Then an H-valued random variable $x(\omega)$ is a map $x : \Omega \to H$ which is measurable relative to μ; furthermore, if $x \in L(\Omega, H)$, then its expectation is given by

$$Ex = \int_\Omega x(\omega)\,d\mu.$$

An H-valued stochastic process is a transformation $x(t,\omega) : [a,b] \times \Omega \to H$ which is measurable in both t and ω. Another important notion is the conditional expectation, $E[x \mid \mathcal{F}]$, of an H-valued variable relative to a subsigma field $\mathcal{F} \subset \mathcal{A}$. $E[x \mid \mathcal{F}]$ is such that

$$\int_C x(\omega)\,d\mu = \int_C E[x \mid \mathcal{F}]\,d\mu$$

for all C in $\mathcal{F}$.

If $\mathcal{F} = \mathcal{A}[y; y \in H]$, that is, if $\mathcal{F}$ is generated by an H-valued random variable y on $(\Omega, \mathcal{A}, \mu)$, then in this case, we write $E[x \mid y]$ for the conditional expectation of x given y, and $E[x \mid y]$ is called the *best global estimate* of x given y.

If $\mathcal{F} = \mathcal{A}[y; y \in H]$ and both x and y are in $L_2(\Omega, H)$ and $L_2(\Omega, \bar{H})$, respectively ($\bar{H}$ is a Hilbert space), then $E[x \mid y]$ is the projection of X on the subspace H_y of $L_2(\Omega, H)$ where

$$H_y = \{u(\omega) \in H : u(\omega) = F(y(\omega)),\ F : \bar{H} \to H\}$$

is measurable relative to the measure induced by y on $(\bar{H}, B(\bar{H}))$.

In the following, we give a brief introduction to the notions of covariance, independence, orthogonality, and Wiener process in Hilbert space (for more details see [15–23]).

The covariance operator of H-random operators x and y is denoted by $\operatorname{cov}(x, y)$ and is given by

$$\operatorname{cov}(x, y) = E(x \cdot y) - (Ex) \cdot (Ey),$$

where $\operatorname{cov}(x, y) \in L(H, H)$ and $(x \cdot y) : H \to H$ defined by $(x \cdot y)h = x\langle y, h\rangle$, for all $h \in H$ with $\langle\ ,\ \rangle$ as the inner product associated with H.

Two H-valued random operators x and y are said to be independent if $E(x \cdot y) = (Ex) \cdot (Ey)$, and to be orthogonal if $E(x \cdot y) = 0$.

An H-valued stochastic process $x(t) = x(t, \omega)$ on $[a, b]$ with $E\|x(t) - x(s)\|^2 < \infty$ is said to have orthogonal (independent) increments if $(x(t_2) - x(s_2))$ and $(x(t_1) - x(s_1))$ are orthogonal (independent), respectively, for all $s_1 < t_1 < s_2 < t_2$ in $[a, b]$. An H-valued stochastic process $W(t) = W(t, \omega)$ on $[a, b]$ is a Wiener process if

(i) $EW(t) = 0$, for all t,
(ii) $W(t)$ is continuous in t with probability one (w.p.1),
(iii) $E(W(t) - W(s) \cdot (W(t) - W(s)) = |t - s|_k$, where $k : H \to H$ is a compact positive definite operator,
(iv) $W(t)$ has independent increments, and
(v) $W(t)$ has orthogonal increments.

If $\phi(t)$, in the second integral of (1), is deterministic and continuously differentiable, then all the integrals in (1) are ordinary Riemann integrals since, in this case, integration by parts can be used for the second integral of (1), that is,

$$\int_a^t \phi(s)\, dW(s) = \phi(t)W(t) - \phi(a)W(a) - \int_a^t \left(\frac{d}{ds}\phi(s)\right) W(s)\, ds,$$

where

$$\int_a^t \left(\frac{d}{ds}\phi(s)\right) W(s)\, ds$$

is an ordinary Riemann integral evaluated for the individual sample functions of $W(s)$.

On the other hand, if $\phi(s)$ is not continuously differentiable or if $\phi(s)$ is an H-valued random transformation, then we cannot, in general, treat the integral

$$\int_a^t \phi(s)\, dW(s)$$

as an ordinary Riemann-Stieltjes integral since almost all sample functions of $W(s)$ are of unbounded variation. Hence, Ito-integral was introduced for a larger class associated with a family $\mathcal{F}_t$ of σ-algebras of $\mathcal{A}$ for t in $[a, b]$ (for details, see [21,24] for the finite-dimensional case and [19] for the infinite-dimensional case).

DEFINITION 1. *Let $u(t) = u(t,\omega)$ be an H-valued stochastic process on $[a,b]$ given by (1), with $u(a) \in L_2(\Omega, H)$, $\phi \in L_2(\Omega, H)$, ϕ measurable relative to $\mathcal{F}_t$ and $A(s) \in L(H)$. Then $u(t)$ is said to have the stochastic differential*

$$du(t) = A(t)u(t)\,dt + \phi(t)\,dW(t). \tag{2}$$

DEFINITION 2. *An H-valued stochastic process $u(t)$ is a solution of (1), and hence, of (2) if*

(i) *$u(t)$ satisfies (1) with probability one,*
(ii) *$u(t)$ is measurable relative to $\mathcal{F}_t$, for all t in $[a,b]$, and*
(iii) *$u(t)$ is continuous with probability one in t.*

Two solutions $u(t)$ and $\bar{u}(t)$ of (2) are the same if $\mu\{\omega \in \Omega : \sup \|u(t) - \bar{u}(t)\| = 0\} = I$, $a < t < b$.

We are now ready to present our estimation problem. We consider the following stochastic dynamical differential equations:

$$\begin{aligned} dX(t) &= A(t)X(t)\,dt + B(t)\,dW(t), \\ X(a,\omega) &= X(a), \qquad E\|X(a)\|^2 < \infty \end{aligned} \tag{3}$$

and

$$\begin{aligned} dY(t) &= C(t)X(t)\,dt + D(t)\,dZ(t), \\ Y(a,\omega) &= 0, \qquad \text{with } a \le t \le b. \end{aligned} \tag{4}$$

Equations (3) and (4) are known in *estimation theory* to generate a *signal process* $X(t) = X(t,\omega)$ and an observation process $Y(t) = Y(t,\omega)$, respectively, where $X(t)$ and $Y(t)$ are H-valued stochastic processes on $[a,b]$ and $X(a)$ is a random Gaussian process in Hilbert space (see [25] for H-valued Gaussian process).

We assume here also that $D(t)$ is such that $D(t)D^*(t)$ is positive definite, and that $A(t)$, $B(t)$, $D(t)$, and $(D(t)D^*(t))^{-1}$ are deterministic bounded operators on H, for t in $[a,b]$. $W(t)$ and $Z(t)$ are H-valued Gaussian Wiener processes which are independent and orthogonal, that is, $\operatorname{cov}(W(t), Z(s)) = 0$ and $E(W(t) \cdot Z(s)) = 0$, respectively, for all t and s in $[a,b]$.

Without loss of generality, we assume that both Wiener processes are normalized, then

$$E\left(W(t) \cdot W^*(s)\right) = \delta(t-s) = E\left(Z(t) \cdot Z^*(s)\right),$$

where δ is the known delta function.

Finally, we assume that

$$\operatorname{cov}(W(t), X(a)) = 0 = \operatorname{cov}(Z(t), X(a)), \qquad \text{for all } t \ge a.$$

For our purpose, the following definition is needed.

DEFINITION 3. *The best linear estimate $\hat{x}$ of $x \in L_2(\Omega, H)$ given $y \in L_2(\Omega, H)$ is $\hat{x} = \Gamma y$, where $\Gamma \in L(\bar{H}, H)$ is such that $E\langle h, (x - \Gamma y)\rangle^2$ minimized for all $h \in H$ and $\Gamma \in L(\bar{H}, H)$.*

If x and y are Gaussians, then it can be shown (see [16]) that the best linear estimate and the best global estimate of x given y are the same, that is, $\hat{x} = E[x \mid y]$.

As in the finite-dimensional case, we define the *innovations process* associated with equation (4) (see [26]) by

$$dn(s) = dY(s) - C(s)\hat{X}(s)\,ds, \tag{5}$$

where $\hat{X}(s)$ is known as the filtered estimate of $X(s)$ and $n(s) = n(s,\omega)$ is a process with orthogonal increments. Now, with all of the above information, we are in a position to state our estimation problem.

3. ESTIMATION PROBLEM

With the standing assumptions on equations (3) and (4) and with the innovations process (5), we state the following.

(i) We wish to find the *best global estimate* of $X(s)$ given the observation process $Y(s)$, with $a < s < t < b$ where t is considered fixed. That is, we want to find the linear estimate of $X(s)$ given by

$$\hat{X}(s \mid t) = \int_a^t K(s,u)\, dn(u),$$

where $K(s,\cdot) \in L(H,H)$ and such that it minimizes

$$E\left\langle h, \left(X(s) - \hat{X}(s \mid t)\right)\right\rangle^2, \qquad \text{for all } h \subset H.$$

If $s < t$, then $\hat{X}(s \mid t)$ is known as the *smoothed estimate* of $X(s)$, and if $s = t$, then $\hat{X}(s \mid s) = \hat{X}(s)$ is the *filtered estimate* of $X(s)$.

(ii) For a given observation $Y(s)$, we wish to use the input-and-output system to establish some relations between the smoothed and filtered estimates of $X(s)$.

4. PROPOSITIONS

Part (i) of the estimation problem depends heavily on the following proposition which we state without proof (see [16,22]), also keeping in mind that since $X(a)$, $W(t)$, and $Y(t)$ are Gaussians, $\hat{X}(s \mid t) = E[X(s) \mid Y_t]$, where $Y_t = \{Y(s) : a \le s \le t\}$.

PROPOSITION 1. *The*

$$\hat{X}(s \mid t) = \int_a^t K(s,u)\, dn(u)$$

is the best linear estimate if and only if

$$E\left(X(s) - \hat{X}(s \mid t) \cdot (n(\sigma) - n(\tau))\right) = 0, \qquad \text{for } a \le \tau \le \sigma < t.$$

Furthermore, $K(s,u) = R(s,u)H(u)$ is the unique solution which satisfies the estimation problem (i), where

$$H(u) = C^*(u)(D(u)D^*(u))^{-1}$$

and

$$R(s,u) = E\left(\left(\hat{X}(s) - X(s)\right) \cdot \left(\hat{X}(u) - X(u)\right)^*\right) = \begin{cases} \phi(s,u)R(u), & s > u, \\ R(s)\phi^*(u,s), & s < u, \end{cases}$$

with $\phi(s,u)$ being the fundamental solution associated with $(A(s) - R(s)H(s)C(s))$ and $R(u) = R(u,u)$ satisfying the following Riccati equation:

$$\frac{d}{du}R(u) = B(u)B^*(u) + A(u)R(u) + R(u)A^*(u) - R(u)H(u)C(u)R(u). \tag{6}$$

From the above proposition, it follows that

$$\hat{X}(s \mid t) = \hat{X}(s) + R(s)\lambda(s \mid t), \tag{7}$$

where

$$\lambda(s \mid t) = \int_s^t \phi^*(u,s)H(u)\, dn(u), \tag{8}$$

with $\lambda(t \mid t) = 0$.

Before proceeding with our approach which uses (7) and (8), we state some known results for the given system (3) and (4) under the standing assumptions.

(I) The filtered estimate $\hat{X}(s)$ satisfies the celebrated Kalman filters equation,

$$d\hat{X}(s) = A(s)\hat{X}(s)\,ds + R(s)H(s)\,dn(s). \tag{9}$$

For details, see [20,27] for the finite-dimensional case and [15,22] for the infinite-dimensional case.

(II) The smoothed estimate $\hat{X}(s \mid t)$ and the *adjoint state* $\lambda(s \mid t)$ satisfy the *Hamiltonian system* given by

$$d\hat{X}(s \mid t) = A(s)\hat{X}(s \mid t)\,ds + B(s)B^*(s)\lambda(s \mid t)\,ds, \tag{10}$$

$$-d\lambda(s \mid t) = -H(s)C(s)\hat{X}(s \mid t)\,ds + A^*(s)\lambda(s \mid t)\,ds + H(s)\,dY(s), \tag{11}$$

where (10) and (11) were derived by Bryson and Frazier [28] and by Jazwinski [23] using variational methods.

(III) The partitioned equations of Lainiotis [29], obtained by him in a classical approach, are given by

$$\hat{X}(s) = \hat{X}_0(s) + \phi_0(a,s)\hat{X}(a \mid s), \tag{12}$$

$$\hat{X}(a,s) = (E - R(a)O_0(a,s))^{-1}\left(\hat{X}(a) + R(a)\lambda_0(a \mid s)\right), \tag{13}$$

$$R(s) = R_0(S) + \phi_0(a,s)R(a)(E - O_0(a,s)R(a))^{-1}\phi^* \circ (a,s), \tag{14}$$

where

(i) $R_0(s)$ is the solution of (6) with $R(a) = 0$,

(ii) $\hat{X}_0(s)$ is the filtered estimate of $X(s)$ when $\hat{X}(a) = 0$ and $R(a) = 0$,

(iii) ϕ_0, O_0, and λ_0 are given, respectively, by

$$\frac{d}{ds}\phi_0(a,s) = (A(s) - R_0(s)H(s)C(s))\phi_0(a,s), \qquad \phi_0(a,a) = E, \tag{15}$$

$$\frac{d}{ds}O_0(a,s) = \phi_0^*(a,s)H(s)C(s)\phi_0(a,s), \qquad O_0(a,a) = 0, \tag{16}$$

$$\lambda_0(a \mid s) = \int_a^s \phi_0^*(u,a)H(u)\,dn_0(u), \quad \text{with } dn_0(u) = dY(u) - C(u)\hat{X}_0(u)\,du.$$

(iv) Kailath and his colleagues' approach (see [30–32]) in deriving Lainiotis equations used Redheffer scattering theory for the Hamiltonian system (10) and (11) with the boundary conditions $\lambda(s \mid s) = 0$ and $\hat{X}(a \mid s) = \hat{X}(a) + R(a)\lambda(a \mid s)$ associated with the scattering picture as in Figure 1, with

$$S_0(a,s) = \begin{pmatrix} t_0(a,s) & \rho_0(a,s) \\ r_0(a,s) & \tau_0(a,s) \end{pmatrix}$$

as the operator associated with the generator,

$$M(s) = \begin{pmatrix} A(s) & B(s)B^*(s) \\ -H(s)C(s) & A^*(s) \end{pmatrix}$$

when $R(a) = 0$, and

$$S_a = \begin{pmatrix} E & R(a) \\ O & E \end{pmatrix}$$

as the constant operator where E is the identity operator, for the boundary layer with the fact that $S(a,s) = S_a^* S$. Hence, Lainiotis and other results follow.

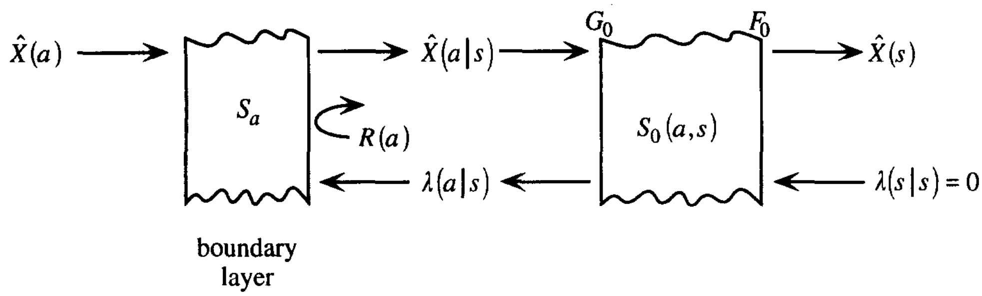


Figure 1. Kailath scattering picture.

Now we are ready to present our results for the estimation problem (ii), but first, we set

$$\tilde{X}(s \mid t) = \hat{X}(s \mid t) - \hat{X}(s) = R(s)\lambda(s \mid t).$$

Then we have the following.

PROPOSITION 2. *$\tilde{X}(s \mid t)$ and $\lambda(s \mid t)$ satisfy the Hamiltonian system given by*

$$d\tilde{X}(s \mid t) = \mathcal{A}(s)\tilde{X}(s \mid t)\,ds + B(s)B^*(s)\lambda(s \mid t)\,ds - R(S)H(s)\,dn(s), \tag{17}$$

$$-d\lambda(s \mid t) = -H(s)C(s)\tilde{X}(s \mid t)\,ds + A^*(s)\lambda(s \mid t)\,ds + H(s)\,dn(S). \tag{18}$$

PROOF. From

$$\lambda(s \mid t) = \int_s^t \phi^*(u,s)H(u)\,dn(u),$$

it follows that

$$\begin{aligned} -\frac{d}{ds}\lambda(s \mid t) &= H(s)\frac{d}{ds}n(s) + \int_t^s \left(\frac{d}{ds}\phi^*(u,s)\right) H(u)\,dn(u), \\ \frac{d}{ds}\phi^*(u,s) &= -A^*(s) - C(s)H(s)R(s)\phi^*(u,s); \end{aligned} \tag{19}$$

hence,

$$\begin{aligned} -\frac{d}{ds}\lambda(s \mid t) &= H(s)\frac{dn}{ds}(s) - (A^*(s) - C(s)H(s)R(s)) \times \int_t^s \phi^*(u,s)H(u)\,dn(u) \\ &= H(s)\frac{dn}{ds}(s) + (A^*(s) - C(s)H(s)R(s))\lambda(s \mid t) \end{aligned}$$

or

$$-d\lambda(s \mid t) = -C(s)H(s)\tilde{X}(s \mid t)\,ds + A(s)\lambda(s \mid t)\,ds + H(s)\,dn(s).$$

On the other hand,

$$\frac{d}{ds}\tilde{X}(s \mid t) = \left(\frac{d}{ds}R(s)\right)\lambda(s \mid t) + R(s)\frac{d}{ds}\lambda(s \mid t)$$

and by (6),

$$\begin{aligned} \frac{d}{ds}\tilde{X}(s \mid t) &= (B(s)B^*(s) + A(s) + R(s)A^*(s) - R(s)H(s)C(s)R(s))\lambda(s \mid t) \\ &\quad + R(s)\left[-H(s)\frac{dn}{ds}(s) - A^*(s) - C(s)H(s)R(s)\lambda(s \mid t)\right] \\ &= A(s)\tilde{X}(s \mid t) + B^*(s)\lambda(s \mid t) - R(s)H(s)\frac{d}{ds}n(s), \end{aligned}$$

and hence, we get (17).

Equations (17) and (18) can be written in matrix form as follows:

$$d\begin{pmatrix} \tilde{X}(s \mid t) \\ -\lambda(s \mid t) \\ E \end{pmatrix} = \begin{pmatrix} A(s) & B(s)B^*(s) & -R(s)H(s)\dfrac{d}{ds}n(s) \\ -H(s)C(s) & A^*(s) & H(s)\dfrac{d}{ds}n(s) \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \tilde{X}(s \mid t) \\ \lambda(s \mid t) \\ E \end{pmatrix} ds.$$

Since $\tilde{X}(s \mid t) = \hat{X}(s \mid t) - \hat{X}(s)$ and $\lambda(s \mid s) = 0$, then (19) becomes

$$\begin{aligned} &d\begin{pmatrix} \hat{X}(s \mid t) \\ -\lambda(s \mid t) \\ E \end{pmatrix} - \begin{pmatrix} A(s) & B(s)B^*(s) & 0 \\ -H(s)C(s) & A(s) & H(s)\dfrac{d}{ds}Y(s) \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{X}(s \mid t) \\ \lambda(s \mid t) \\ E \end{pmatrix} ds \\ &= d\begin{pmatrix} \hat{X}(s) \\ -\lambda(s \mid s) \\ E \end{pmatrix} - \begin{pmatrix} A(s) & B(s)B^*(s) & R(s)H(s)\dfrac{d}{ds}n(s) \\ -H(s)C(s) & A^*(s) & H(s)\dfrac{d}{ds}(Y(s) - n(s)) \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{X}(s) \\ \lambda(s \mid s) \\ E \end{pmatrix} ds. \end{aligned} \tag{20}$$

From the input-and-output point of view, if we consider $\tilde{X}(s \mid t)$ and $\lambda(s \mid t)$ as waves through a slab located between a and t, where t is fixed, then from $\tilde{X}(s \mid t) = \hat{X}(s \mid t) - \hat{X}(s)$, $\lambda(s \mid t) = \lambda(s \mid t) - \lambda(s \mid s)$ then physically the waves $X(s \mid t)$ and $\lambda(s \mid t)$ can be decomposed in internal and external states, that is,

$$\begin{pmatrix} \hat{X}(s \mid t) \\ -\lambda(s \mid t) \\ E \end{pmatrix}$$

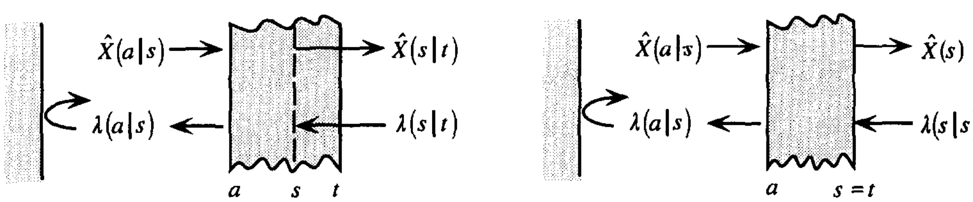
will be considered as the internal state of

$$\begin{pmatrix} \tilde{X}(s \mid t) \\ -\lambda(s \mid t) \\ E \end{pmatrix}$$

and

$$\begin{pmatrix} \hat{X} \\ -\lambda(s \mid s) \\ E \end{pmatrix}$$

as the external state, as seen in Figure 2.



Internal state External state

Figure 2.

The decomposition of the waves $X(s \mid t)$ and $\lambda(s \mid t)$ into internal state and external state along with (20) suggests the decomposition of (19) into internal and external state equations, respectively, by

$$d\begin{pmatrix} \hat{X}(s \mid t) \\ -\lambda(s \mid t) \\ E \end{pmatrix} = N_1(s) \begin{pmatrix} \hat{X}(s \mid t) \\ -\lambda(s \mid t) \\ E \end{pmatrix} ds, \qquad a \le s < t \quad (t \text{ fixed}), \tag{21}$$

where $N_1(s)$ is the matrix on the left-hand side of (20) and

$$d\begin{pmatrix} \hat{X}(s) \\ -\lambda(s \mid s) \\ E \end{pmatrix} = N_E(s)\begin{pmatrix} \hat{X}(s) \\ \lambda(s \mid s) \\ E \end{pmatrix} ds, \qquad a \le s = t, \tag{22}$$

where $N_E(s)$ is the right-hand side matrix of (20).

PROPOSITION 3. *The decomposition of system (19) into internal state equation (21) and external state equation (22) is necessary and sufficient.*

PROOF. The proof of sufficiency is trivial. To prove the necessity, we note that since $dn(s) = dY(s) - C(s)\hat{X}(s)\,ds$, then equation (18)

$$-d\lambda(s \mid t) = -H(s)C(s)\hat{X}(s \mid t)\,ds - \hat{X}(s)\,ds$$

becomes

$$A^*(s)\lambda(s \mid t)\,ds + H(s)\,dy(s) - C(s)\hat{X}(s)\,ds$$
$$= -H(s)C(s)\hat{X}(s \mid t)\,ds + A^*(s)\lambda(s \mid t)\,ds + H(s)\,dy(s),$$

and hence, we get the equation of $\lambda(s \mid t)$ in (21). $\lambda(s \mid s)$ in (22) is trivial since $\lambda(s \mid s) = 0$.

Now since $\hat{X}(s \mid t) = \hat{X}(s) + R(s)\lambda(s \mid t)$, then

$$\begin{aligned} d\hat{X}(s \mid t) &= d\hat{X}(s) + \left(\frac{d}{ds}R(s)\right)\lambda(s \mid t)\,ds + R(s)\,d\lambda(s \mid t) \\ &= d\hat{X}(s) + \left(\frac{d}{ds}R(s)\right)\lambda(s \mid t)\,ds + R(s)H(s)C(s)\hat{X}(s \mid t)\,ds \\ &\quad - A^*(s)\lambda(s \mid t)\,ds - H(s)\,dY(s), \end{aligned}$$

and hence,

$$d\hat{X}(s \mid t)\Big|_{t=s} = d\hat{X}(s) - R(s)H(s)\,dn(s).$$

On the other hand, we have from (21)

$$d\hat{X}(s \mid t)\Big|_{t=s} = A(s)\hat{X}(s)\,ds;$$

hence,

$$d\hat{X}(s) - R(s)H(s)\,dn(s) = A(s)\hat{X}(s)\,ds$$

and the last equation is the equation of $\hat{X}(s)$ in (22).

LEMMA 1. *The internal state equation (21) gives the Hamiltonian system (10),(11) for the smoothed estimate $\hat{X}(s \mid t)$ and the adjoint state $\lambda(s \mid t)$. The external state equation (22) gives the Kalman filter equation (9) for the filtered estimate $\hat{X}(s)$. The proof is straightforward and is omitted; when we examine the internal state (21), that is,*

$$N_1(s) = \begin{pmatrix} A(s) & B(s)B^*(s) & 0 \\ -H(s)C(s) & A^*(s) & H(s)\dfrac{dY}{ds}(s) \\ 0 & 0 & 0 \end{pmatrix}$$

has an operator

$$P_0(a,s) = \begin{pmatrix} t_0(a,s) & \rho_0(a,s) & F_0(a,s) \\ r_0(s,a) & t_0^*(a,s) & G_0(a,s) \\ 0 & 0 & E \end{pmatrix},$$

when $R(a) = 0$.

From [33], it can be shown that

$$\frac{d}{ds}t_0(a,s) = (A(s) - \rho_0(a,s)H(s)C(s))t_0(a,s), \qquad t_0(a,a) = E, \tag{23}$$

$$\begin{aligned}\frac{d}{ds}\rho_0(a,s) &= B(s)B^*(s) + A(s)\rho_0(a,s) + \rho_0(a,s)A^*(s) \\ &\quad - \rho_0(a,s)H(s)C(s)\rho_0(a,s), \qquad \rho_0(a,a) = 0,\end{aligned} \tag{24}$$

$$\frac{d}{ds}r_0(a,s) = t_0^*(a,s)H(s)C(s)t_0(a,s), \qquad r_0(a,a) = 0, \tag{25}$$

$$\begin{aligned}dF_0(a,s) &= (A(s) - \rho_0(a,s)H(s)C(s))F_0(a,s)\,ds \\ &\quad + \rho_0(a,s)H(s)\,dY(s), \qquad F_0(a,a) = 0,\end{aligned} \tag{26}$$

$$dG_0(a,s) = t_0^*(a,s)H(s)(dY(s) - C(s)F_0(a,s)\,ds), \qquad G_0(a,a) = 0. \tag{27}$$

We note that equation (24) is the same as equation (6) with $\rho_0(a,a) = 0 = R(a)$, hence, $R(s) = \rho_0(a,s)$. Also, equation (23) is the same as equation (15) and this gives $\phi_0(a,s) = t_0(a,s)$ and equations (25) and (16) imply that $O_0(a,s) = r_0(a,s)$.

Now if

$$P(a,s) = \begin{pmatrix} t(a,s) & \rho(a,s) & F(a,s) \\ r(a,s) & t^*(a,s) & G(a,s) \\ O & O & E \end{pmatrix}$$

is the operator associated with the generator $N_1(s)$ when $\rho(a,a) = R(a) = 0$, $t(a,a) = t_a$, $r(a,a) = 0$, $F(a,a) = F_a$, and $G(a,a) = G_a$, then $P(a,s) = P_a * P_0(a,s)$, see [34] for details, where

$$P_a = \begin{pmatrix} t_a & R(a) & F_a \\ O & t_a^* & G_a \\ O & O & E \end{pmatrix}$$

is the modified initial condition matrix, and provided $(E - R(a)r_0(a,s))^{-1}$ exists, in which case we have

$$P_{(a,s)} = \begin{pmatrix} t_0(a,s)Pt_a & \rho_a(a,s)+t_0(a,s)R(a)qt_0^*(a,s) & t_0(a,s)PF_a+t_0(a,s)R(a)qG_0(a,s)+F_0(a,s) \\ t_a^*r_0(a,s)Pt_a & t_a^*qt_0^*(a,s) & t_a^*r_0(a,s)PF_a+t_a^*qG_0(a,s)+G_a \\ O & O & E \end{pmatrix}, \tag{28}$$

with $P = (E - R(a)r_0(a,s))^{-1}$ and $q = (E - r_0(a,s)R(a))^{-1}$.

The elements of $P(a,s)$ satisfy the same differential equations as those of $P_0(a,s)$; in particular, we need

$$dF(a,s) = (A(s) - \varphi(a,s)H(s)C(s))F(a,s)\,dS + \varphi(a,s)H(s)\,dY(s), \qquad F(a,a) = F_a, \tag{29}$$

$$dG(a,s) = t^*(a,s)H(s)(dY(s) - C(s)F(a,s)\,dS), \qquad G(a,a) = G_a. \tag{30}$$

We consider the internal state (21) as pictured in Figure 3 with an arbitrary boundary condition. Then

$$\begin{pmatrix} \hat{X}(s \mid t) \\ \lambda(a \mid s) \\ E \end{pmatrix} = P(a,s) \begin{pmatrix} \hat{X}(a \mid s) + \gamma \\ r\lambda(s \mid t) \\ E \end{pmatrix}, \qquad a \le s < t,$$

or

$$\hat{X}(s \mid t) = t(a,s)\left(\hat{X}(a \mid s) + \gamma\right) + \rho(a,s)\lambda(a \mid t) + F(a,s), \tag{31}$$

$$\lambda(a \mid s) = r(a,s)\left(\hat{X}(a \mid s) + \gamma\right) + t^*(a,s)\lambda(s \mid t) + G(a,s). \tag{32}$$

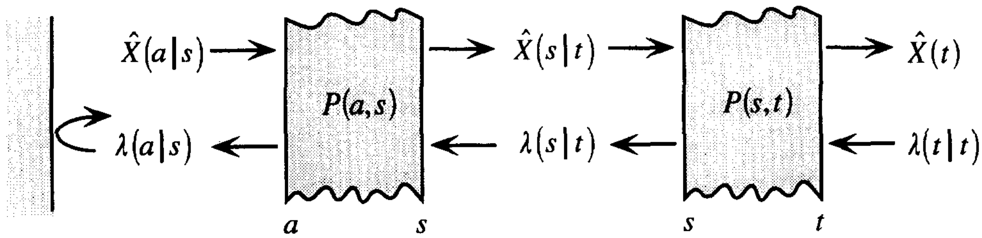


Figure 3. Internal state.

LEMMA 2. *The internal state given by (31) and (32) with zero inputs implies that*

$$\hat{X}(s) = F(a,s),$$
$$\lambda(a \mid s) = G(a,s).$$

PROOF. Zero inputs means that $\hat{X}(a \mid s) = 0$ and $\lambda(s \mid t) = 0$ as $s \to t$ and $\gamma = 0$, and the proof then becomes trivial.

If we denote the solutions of the internal state with zero inputs by $\hat{\mathbf{X}}(s)$ and $\boldsymbol{\lambda}(a \mid s)$, then as an immediate consequence of the previous lemma, which satisfies the Kalman filters equation since equation (3) gives

$$d\hat{\mathbf{X}}(s) = dF(a,s) = (A(s) - \rho(a,s)H(s))C(s)\hat{\mathbf{X}}(s)\,ds + \rho(a,s)H(s)\,dY(s), \qquad \hat{X}(a) = F_a$$

and that equation (3) also gives

$$d\boldsymbol{\lambda}(a \mid s) = dG(a,s) = t^*(a,s)H(s)(dY(s) - C(s)\hat{\mathbf{X}}(s)\,ds), \qquad \boldsymbol{\lambda}(a \mid a) = G_a.$$

PROPOSITION 4. *The internal state given by (31) and (32) with $s \to t$ gives*

$$\hat{X}(s) = t(a,s)\left(\hat{X}(a \mid s) + \gamma\right) + \hat{\mathbf{X}}(s), \tag{33}$$
$$\lambda(a \mid s) = r(a,s)\left(\hat{X}(a \mid s) + \gamma\right)\boldsymbol{\lambda}(a \mid s), \tag{34}$$
$$\hat{X}(a \mid s) = E - R(a)(r(a,s))^{-1}[X(a) + R(a)(\boldsymbol{\lambda}(a \mid s)) + r(a,s)\gamma]. \tag{35}$$

PROOF. Since $\hat{X}(s)$ and $\boldsymbol{\lambda}(a \mid s) = G(a,s)m$, then with $s \to t$ (33) and (34) follow immediately from (31) and (32), respectively.

To establish (35), we get from $\hat{X}(a \mid s) = \hat{X}(a) + R(a)\lambda(a \mid s)$ and from (34) the following:

$$\hat{X}(a \mid s) = \hat{X}(a) + R(a)[r(a,s)(X(a \mid s) + \gamma) + \boldsymbol{\lambda}(a \mid s)]$$

and

$$\hat{X}(a \mid s) = (E - R(a)r(a,s))^{-1}\left[\hat{X}(a) + R(a)(\boldsymbol{\lambda}(a \mid s) + r(a,s)\gamma)\right],$$

provided that $(E - R(a)r(a,s))^{-1}$ exists.

We should note that we can have a zero input $\hat{X}(a \mid s)$ by having $\hat{X}(a) = 0$ and $R_0(a) = 0$. In this case, we denote the solution of the dynamic system, given by (3) and (4), by $X_0(s)$ and the adjoint state by $\lambda_0(s \mid t)$, and if $t_0 = E$, $\gamma = F_a = G_a = 0$, then we get the Kailath and Lainiotis results mentioned earlier, that is,

$$\hat{X}_0(s) = \hat{X}(s) = F(a,s) = F_0(a,s),$$
$$\lambda_0(a \mid s) = \lambda(a \mid s) = G(a \mid s) = G_0(a,s),$$

and equations (33) and (35) become, respectively,

$$\hat{X}(s) = t_0(a,s)\hat{X}(a \mid s) + \hat{X}_0(s)$$

and

$$\hat{X}(a \mid s) = (E - R(a)r_0(a,s))^{-1}\left(\hat{X}(a) + R(a)\lambda_0(a \mid s)\right),$$

which are Lainiotis equations (12) and (13), respectively, and Lainiotis equation (14), namely,

$$\rho(a,s) = \rho_0(a,s) + t_0(a,s)R(a)(E - r_0(a,s)R(a))^{-1}t_0(a,s)$$

can be obtained directly from (28).

Now we examine the external state given by (22) where the generator $N_E(s)$ can be rewritten as

$$\mathbf{N}_E(s) = \begin{pmatrix} A(s) - R(s)H(s)C(s) & B(s)B^*(s) & R(s)H(s)\dfrac{dY}{ds} \\ O & A^*(s) & O \\ O & O & O \end{pmatrix}$$

and we assume that the generator $\mathbf{N}_E(s)$ has an operator $P_0(a,s)$ when $R_0(a,s) = 0$ with

$$N_E(s) = \begin{pmatrix} \mathbf{t}_0(a,s) & \rho_0(a,s) & \mathbf{F}_0(a,s) \\ \mathbf{r}_0(a,s) & \tau_0(a,s) & \mathbf{G}_0(a,s) \\ O & O & E \end{pmatrix}, \qquad a \le s = t.$$

It can be shown, see [33], that

$$\frac{d\mathbf{t}_0}{ds}(a,s) = (A(s) - R_0(s)H(s)C(s))\mathbf{t}_0(a,s), \qquad t_0(a,a) = E,$$

hence, $\phi(a,s) = t_0(a,s) = \mathbf{t}_0(a,s)$,

$$\frac{d\rho_0}{ds}(a,s) = B(s)B^*(s) + A(s) - R_0(s)H(s)C(s)\rho_0(a,s) + \rho_0(a,s)A^*(s), \qquad \rho_0(a,a) = 0.$$

But this implies that we must have $\rho_0(a,s) = R_0(s)$, $d\mathbf{r}_0/ds(a,s) = 0$, $\mathbf{r}_0(a,a) = 0$, hence, $\mathbf{r}_0(a,s) = 0$, $d\tau_0/ds(a,s) = \tau_0(a,s)A^*(s)$, $\tau_0(a,a) = E$, $d\mathbf{F}_0(a,s) = (A(s) - R_0(s)H(s)C(s))$ $\mathbf{F}_0(a,s)\,ds + R_0(s)H(s)\,dY(s)$, $\mathbf{F}_0(a,a) = 0$, hence, $uF_0(a,s) = F_0(a,s)$, $d\mathbf{G}_0(a,s) = 0$, $\mathbf{G}_0(a,a) = 0$, hence, $\mathbf{G}_0(a,s) = 0$. In the case where the operator $\mathbf{P}(a,s)$ corresponds to the generator $N_E(s)$ when $R(a) \neq 0$, we have

$$\mathbf{P}(a,s) = \mathbf{P}_a * \mathbf{P}_0(a,s),$$

where

$$\mathbf{P}_a = \begin{pmatrix} \mathbf{t}_a & R(a) & \mathbf{F}_a \\ O & \tau_a & O \\ O & O & E \end{pmatrix}$$

is the modified initial condition,

$$\mathbf{P}(a,s) = \begin{pmatrix} \mathbf{t}_0(a,s)\mathbf{t}_a & \rho_0(a,s) + \mathbf{t}_0(a,s)R(a)\tau_0(a,s) & \tau_0(a,s)\mathbf{F}_a + \mathbf{F}_0(a,s) \\ 0 & \tau_0\tau_0(a,s) & 0 \\ 0 & 0 & E \end{pmatrix}.$$

We note also that the elements of $\mathbf{P}(a,s)$ satisfy the same differential equations as those of $\mathbf{P}_0(a,s)$, and in particular we need

$$d\mathbf{F}(a,s) = (A(s) - R(s)H(s)C(s))\mathbf{F}(a,s)\,ds + R(s)H(s)\,dY(s), \qquad \mathbf{F}(a,a) = \mathbf{F}_a. \tag{36}$$

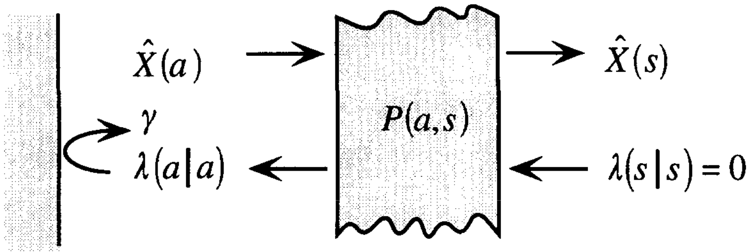


Figure 4. External state $a \le s = t$.

The external state (22) with arbitrary boundary conditions can be pictured as in Figure 4, that is, the external state is equivalent to

$$\begin{pmatrix} \hat{X}(s) \\ \lambda(a \mid a) \\ E \end{pmatrix} = \mathbf{P}(a,s) \begin{pmatrix} \hat{X}(a \mid s) + \gamma \\ \lambda(s \mid t) \\ E \end{pmatrix}$$

or

$$\hat{X}(s) = \mathbf{t}(a,s)\left(\hat{X}(a) + \gamma\right)\mathbf{F}(a,s). \tag{37}$$

LEMMA 3. *The external state given by (37) with zero inputs implies that*

$$\hat{X}(s) = \mathbf{F}(a,s).$$

The proof is trivial since for the zero inputs we need only to have

$$\hat{X}(a) = 0 \qquad \text{and} \qquad \gamma = 0.$$

Now if we denote the solution of the external state (37) with zero inputs by $\hat{X}(s)$, then (37) becomes

$$\hat{X}(s) = t(a,s)\hat{X}(a) + \hat{X}(s); \tag{38}$$

furthermore, $\hat{X}(s)$ satisfies the Kalman filter equation since (36) gives

$$d\hat{X}(s) = d\mathbf{F}(a,s) = (A(s) - R(s)H(s)C(s))X(s)\,ds + R(s)H(s)\,dY(s), \qquad \hat{X}(a) = \mathbf{F}a.$$

If $\mathbf{F}_a = 0$ and $\mathbf{t}_a = E$, then $\mathbf{t}(a,s) = \mathbf{t}_0(a,s)$ and $\hat{X}_0(s) = \mathbf{F}(a,s) = \mathbf{F}_0(a,s)$, and then (38) becomes

$$\hat{X}(s) = t_0(a,s)\hat{X}(a) + \hat{X}_0(s),$$

but this is just the Lainiotis equation (12) with $\lambda(a \mid s) = 0$ and $\hat{X}(a \mid s) = \hat{X}(a) + R(a)\lambda(a \mid s) = \hat{X}(a)$.

REFERENCES

1. S. Chandrasekhar, *Radiative Transfer*, Dover Publications, (1960).
2. H. Natsuyama, S. Ueno and A. Wang, *Terrestrial Radiative Transfer*, Springer, New York, (1998).
3. R. Redheffer, On the relation of transmission-line theory to scattering and transfer, *J. of Mathematics and Physics* **XLI** (1), 1–41, (1962).
4. V.V. Sobolev, *Light Scattering in Planetary Atmospheres*, (Translated by W.M. Irvine), Pergamon Press, Oxford, UK, (1973).
5. G.M. Wing, *An Introduction to Transport Theory*, Wiley, New York, (1962).
6. R. Redheffer, Microwave antenna and dielectric surface, *J. Appl. Phys.* **20**, 397–411, (1949).
7. A.P. Wang, Invariant imbedding and scattering theory, *J. Math. Appl.* **17**, 48–60, (1967).
8. S. Shaw and A. Wang, Scattering operators and propagation solutions for a non-stationary transfer equation, *J. Math. Analysis Appl.* **134** (1), 170–185, (1988).
9. A.P. Wang, Nonstationary multiple scattering, *J. Math. Phys.* **18**, 47–51, (1977).

10. W.R. Reid, *Riccati Differential Equations*, Academic Press, New York, (1972).
11. R. Bellman, H. Kagiwada, R. Kalaba and S. Ueno, Invariant imbedding and the computation of internal field for transport processes, *J. Math. Anal. Appl.* **12**, 541–548, (1965).
12. D. Dudley and A.P. Wang, Operator theory on WKB method and Bremmer series, *J. Math. Physics* **24** (6), 1470–1476, (1983).
13. G.G. Stokes, On the intensity of the light reflected from or transmitted through a pile of plates, *Proc. Roy. Soc.* **11**, 545–556, (1862).
14. A.P. Wang, An inverse problem of reflection, *Computers Math. Applic.* **25** (9), 59–65, (1993).
15. R.F. Curtain, Estimation theory for abstract evolution equations excited by general white noise processes, *SIAM J. Control and Optimization* **14**, 1124–1150, (1976).
16. R.F. Curtain, Infinite dimensional estimation theory for linear systems, Control Theory Cent. Rep. No. 38, University of Warwick, Coventry, England, (1976).
17. R.F. Curtain, Infinite-dimensional filtering, *SIAM J. Control* **13**, 89–103, (1975).
18. R.F. Curtain, Stochastic differential equations in Hilbert space, Ph.D. Dissertation, Browman University, Providence, RI, (1969).
19. R.F. Curtain and P.L. Falb, Stochastic differential equations in Hilbert space, *J. Differential Equations* **10**, 412–430, (1971).
20. M.H.A. Davis, *Linear Estimation and Stochastic Control*, John Wiley, New York, (1977).
21. J.L. Doob, *Stochastic Processes*, John Wiley, New York, (1953).
22. P.L. Falb, Infinite dimensional filtering: The Kalman-Bucy filter in Hilbert space, *Information and Control* **11**, 102–137, (1967).
23. A.H. Jazwinski, *Stochastic Processes and Filtering Theory*, Academic Press, New York, (1970).
24. L. Arnold, *Stochastic Differential Equations: Theory and Applications*, John Wiley, New York, (1973).
25. A.T. Bharucha-Reid, *Random Integral Equations*, Academic Press, New York, (1972).
26. T. Kailath, An innovations approach to least-square estimation. Part I: Linear filtering in additive white noise. Part II: Smoothing in additive white noise (with P. Frost), *IEEE, Trans. Automatic Control* **AC-13**, 646–660, (1968).
27. R. Bucy and P.D. Joseph, *Filtering for Stochastic Processes with Applications to Guidance*, Interscience, New York, (1968).
28. A.E. Bryson and M. Frazier, Smoothing for linear and nonlinear dynamic systems, TDR 63-119, pp. 353–364, Aero. Sys. Div., Wright-Patterson Air Force Base, Dayton, OH, (1963).
29. D.G. Lainiotis, Partitioned estimation algorithms, II: Linear estimation, *Information Science* **7**, 317–340, (1974).
30. T. Kailath, Redheffer scattering theory and linear state space estimation problems, *Ricerche di Automatica* **10**, 135–163, (1979).
31. T. Kailath and L. Ljung, *A Scattering Theory Frame Work for Fast Least-Squares Algorithms, Multivariant Analysis VI*, pp. 387-406, North-Holland, Amsterdam, (1977).
32. L. Ljung, T. Kailath and B. Friedlander, Scattering theory and least squares estimation, Part I: Continuous-time problem, *Proc. IEEE* **64**, 131–138, (1976).
33. F. Mistiri, Star-product and linear dynamical systems, Ph.D. Dissertation, Arizona State University, (1983).
34. F. Mistiri and A.P. Wang, The star-product and its algebraic propertie, *J. of the Franklin Institute* **321**, 21–38, (1986).